

5.1 ПОНЯТТЯ ПРОЦЕСУ

Найважливішою функцією ОС є організація раціонального використання усіх її апаратних та інформаційних ресурсів. До основних ресурсів можуть бути віднесені процесори, пам'ять, зовнішні пристрої, дані і програми. Маючи в розпорядженні одні і ті ж апаратні ресурси, але керована різними ОС, обчислювальна система може працювати з різним ступенем ефективності. Тому знання внутрішніх механізмів операційної системи дозволяє побічно судити про її експлуатаційні можливості і характеристики. Хоча і в однопрограми ОС необхідно розв'язувати задачі управління ресурсами, головні складнощі на цьому шляху виникають в мультипрограми ОС, в яких за ресурси конкурують відразу декілька додатків. Саме тому велика частина всіх проблем, що розглядаються тут, стосується також і мультипрограми систем.

Одним з основних понять, пов'язаних з ОС, є процес – абстрактне поняття, що описує роботу програми. Усе програмне забезпечення, що функціонує на комп'ютері, включаючи і ОС, можна представити набором процесів.

Щоб підтримувати мультипрограмування, ОС повинна визначити і оформити для себе ті внутрішні одиниці роботи, між якими розподілятиметься процесор та інші ресурси комп'ютера. Пояснюючи поняття «Операційна система» і описуючи способи побудови ОС, ми часто застосовували слова «програма» і «завдання». Нині в більшості ОС визначені два типи одиниць роботи: більша одиниця – процес, або задача, і менш велика – потік, або нитка. Причому процес виконується у формі одного або декількох потоків.

При використанні цих термінів часто виникають складнощі. Це відбувається в силу декількох причин. По-перше, – специфіка різних ОС, коли співпадаючі по суті поняття одержали різні назви, наприклад задача (task) в OS/2, OS/360 і процес (process) в UNIX, Windows NT, NetWare. По-друге, у міру розвитку системного програмування і методів організації обчислень деякі з цих термінів отримали нове смислове значення, особливо це стосується поняття «процес», який поступився багатьма своїми властивостями новому поняттю «потік». По-третє, термінологічні складнощі породжуються наявністю декількох варіантів перекладу англійських термінів. Наприклад, термін «thread» перекладається як «нитка», «потік», «полегшений процес»,

«мінізадача» тощо. Далі в якості назви одиниць роботи ОС використовуватимуться терміни «процес» і «потік». У тих же випадках, коли відмінності між цими поняттями не гратимуть істотної ролі, вони об'єднуються під узагальненим терміном «задача».

Раніше говорили: обчислювальна система виконує одну або декілька програм, операційна система планує завдання, програми можуть обмінюватися даними тощо. Використовували ці терміни в деякому загальноживаному, життєвому сенсі, припускаючи, що усі читачі однаково уявляють собі, що мається на увазі під ними в кожному конкретному випадку. При цьому одні і ті ж слова означали і об'єкти в статичному стані, що не обробляються обчислювальною системою (наприклад, сукупність файлів на диску), і об'єкти в динамічному стані, що знаходяться в процесі виконання. Це було можливо, поки ми говорили про загальні властивості операційних систем, не вдаючись до подробиць їх внутрішнього устрою і поведінки, або про роботу обчислювальних систем 1-2-го покоління, які не могли обробляти більше однієї програми або одного завдання одночасно, по суті справи не маючи операційних систем. Але тепер ми знайомимося з деталями функціонування сучасних комп'ютерних систем, і нам доведеться уточнити термінологію.

Розглянемо такий приклад. Два студенти запускають програму витягання квадратного кореня. Один хоче обчислити квадратний корінь з 4, а другий – з 9. З точки зору студентів, запущена одна і та ж програма; з точки зору комп'ютерної системи, їй доводиться займатися двома різними обчислювальними процесами, оскільки різні початкові дані призводять до різного набору обчислень. Отже, на рівні того, що відбувається всередині обчислювальної системи ми не можемо використати термін «програма» в призначеному для користувача значенні слова. Розглядаючи системи пакетної обробки, ми ввели поняття «завдання» як сукупність програми, набору команд мови управління завданнями, необхідних для її виконання, і вхідних даних. З точки зору студентів, вони, підставивши різні початкові дані, сформувавши два різні завдання. Можливо, термін «завдання» підійде нам для опису внутрішнього функціонування комп'ютерних систем? Щоб з'ясувати це, давайте розглянемо інший приклад. Нехай знову обидва студенти намагаються витягти корінь квадратний з 9, тобто нехай вони сформувавши ідентичні завдання, але завантажили їх в обчислювальну систему із зрушенням за часом. Тоді як одне з виконуваних завдань приступило до

друку отриманого значення, і чекає закінчення операції введення-виведення, друге тільки починає виконуватися. Чи можна говорити про ідентичність завдань усередині обчислювальної системи в даний момент? Ні, оскільки стан процесу їх виконання різний. Отже, і слово «завдання» в призначеному для користувача сенсі не може застосовуватися для опису того, що відбувається в обчислювальній системі.

Це відбувається тому, що терміни «програма» і «завдання» призначені для опису статичних, неактивних об'єктів. Програма ж у процесі виконання є динамічним, активним об'єктом. По ходу її роботи комп'ютер обробляє різні команди і перетворює значення змінних. Для виконання програми ОС повинна виділити певну кількість оперативної пам'яті, закріпити за нею певні пристрої введення-виведення або файли, тобто зарезервувати певні ресурси із загального числа ресурсів усієї обчислювальної системи. І, звичайно ж, неможливе виконання програми без надання їй процесорного часу, тобто часу, впродовж якого процесор виконує коди цієї програми.

Кількість ресурсів і конфігурація системи з часом можуть змінюватися. Тому для опису таких активних об'єктів усередині комп'ютерної системи замість термінів «програма» і «завдання» використовується термін – «процес». Таким чином, процес – це контейнер, в якому міститься вся інформація, необхідна для роботи програми [9]. Термін «процес» уперше почали застосовувати розробники системи MULTICS в 1960-х роках.

Процес знаходиться під управлінням ОС, тому в ньому може виконуватися не лише код програми, але і частина коду ядра ОС (що не знаходиться у виконуваному файлі), як у випадках, спеціально запланованих авторами програми (наприклад, при використанні системних викликів), так і в непередбачених ситуаціях (наприклад, при обробці зовнішніх переривань). В операційних системах, де існують і процеси, і потоки, процес розглядається ОС як заявка на споживання усіх видів ресурсів, окрім одного – процесорного часу. Цей останній найважливіший ресурс розподіляється ОС між іншими одиницями роботи – потоками, які і одержали свою назву завдяки тому, що вони є послідовністю (потоки виконання) команд.

У простому випадку процес складається з одного потоку, і саме таким чином трактували поняття «процес» до середини 1980-х років (наприклад, в ранніх версіях UNIX) і в такому ж виді він зберігся в деяких сучасних ОС. У таких системах поняття

«потік» повністю поглинається поняттям «процес», тобто залишається тільки одна одиниця роботи і споживання ресурсів – процес. Мультипрограмування здійснюється в таких ОС на рівні процесів. Відмітимо, що в однієї програмних системах не виникає необхідності введення поняття, що означає одиницю роботи, оскільки там не існує проблеми розподілу ресурсів. Взаємозв'язок між завданнями, процесами і потоками показаний на рис. 5.1.

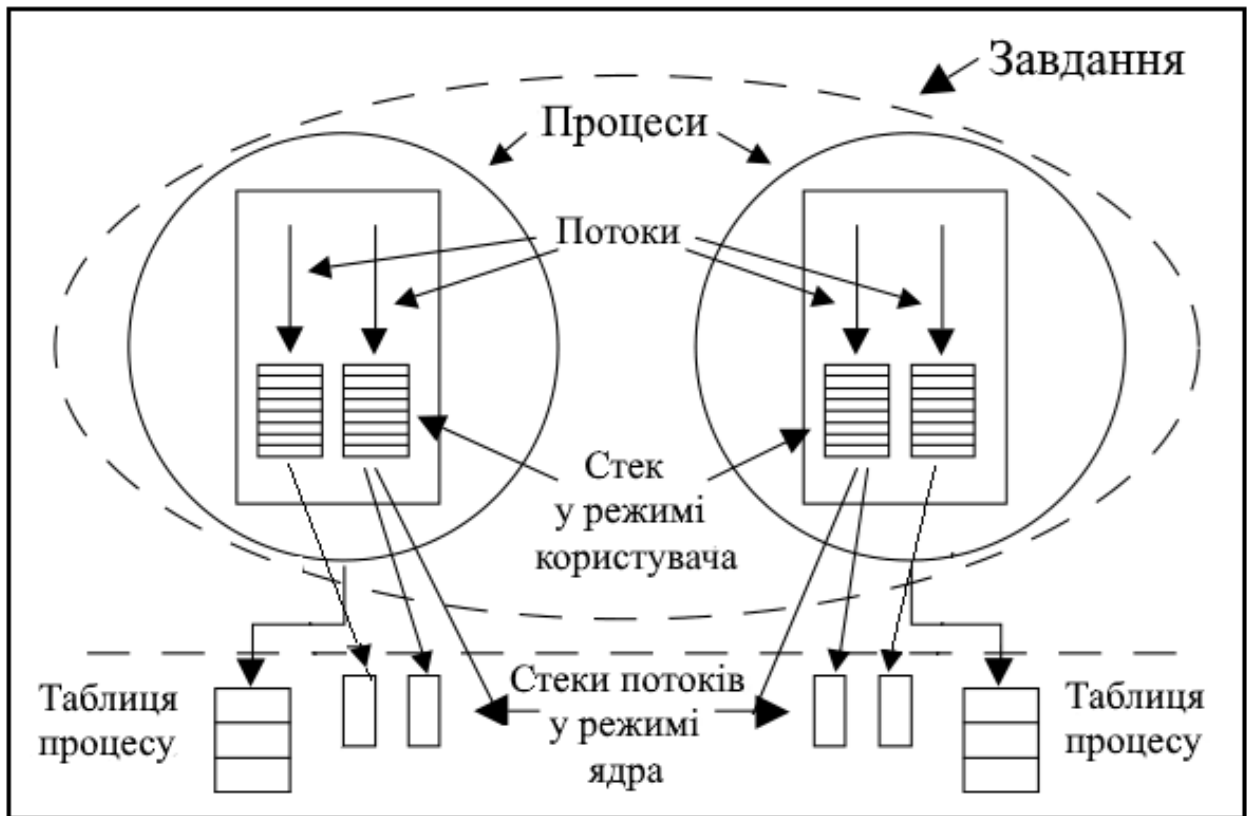


Рисунок 5.1 – Завдання, процеси, потоки

Таким чином, під процесом розуміють програму разом з даними і необхідними ресурсами, яка виконується обчислювальною системою, включаючи поточні значення лічильника команд і регістрів. Тобто, усе програмне забезпечення, у тому числі ОС, які виконуються на комп'ютері, є рядом процесів.